

Personalized Telerobotics by Fast Machine Learning of Body-Machine Interfaces

Matteo Macchini, Fabrizio Schiano, and Dario Floreano

Abstract—Human-Robot Interfaces (HRIs) can be hard to master for inexperienced users, making the teleoperation of mobile robots a difficult task. The development of Body-Machine Interfaces (BoMIs) represents a promising approach to making a user more proficient, by exploiting the natural control they can exert on their own body motion. Since human motion presents individual traits due to several factors, including physical condition, age, and experience, generic BoMIs still require a significant learning time and effort to reach adequate ability in teleoperation. In this work, we present a novel approach which provides a Body-Machine Interface tailored on the specific user. Our method autonomously learns from the user their preferred strategy to control the robot, and provide a personalized body-machine mapping. We show that the proposed method can significantly reduce the duration of the training phase in teleoperation, thus allowing faster skill acquisition. We validated our approach by performing both simulation and real-world experiments with human subjects. The first involved the teleoperation of a fixed-wing simulated drone, while the second consisted in controlling a real quadrotor. We used our framework to extrapolate common and peculiar features of movements among individuals. Observing reoccurring strategies, we provide insights on how humans would naturally interface with a distal machine.

Index terms — Telerobotics and Teleoperation, AI-Based Methods, Learning and Adaptive Systems, Wearable Robots

Supplementary video: <https://youtu.be/ssLa75fly2Y>

I. INTRODUCTION

The term *telerobotics*, derived from the greek *tele* (distant), refers to a system with a human-in-the-loop operator, controlling a robot situated in a separate environment [1].

Telerobotics finds relevant uses in fields where human cognition and decision-making capabilities cannot be substituted by machine intelligence, including exploration, supervision and maintenance in risky environments, such as nuclear plants [2] or during search and rescue missions [3]. Robotic systems can also be employed in tasks where augmenting the operator's perception and accuracy is needed, as in the case of minimally invasive surgery [4].

These applications require the design and implementation of control interfaces that are sufficiently powerful and intuitive



Fig. 1: Human user (A) teleoperates a quadrotor (B) in first-person view (C) by body motion.

also for inexperienced users. A telerobotic system is defined as a dual architecture where the controlled element (the robot) executes commands issued by the operator through a certain control device [1]. The control device normally consists of a hardware interface such as a joystick, or a steering wheel.

A subdomain of HRIs is represented by body-machine interfaces (BoMIs), which collect information from the human body, and translate them into commands for an external device [5]. Body movements have been largely employed as a source of information for teleoperation as they find foundation in the intuitiveness of a person's control over their own body motion. Here, the control device of the telerobotic system is the user's body itself. When robot and control device are kinematically similar, the HRI is said to be "homologous" [6]. There are several examples in literature on the control of complex systems as humanoid robots or human-like robotic limbs from human movement [7], [8], [9]. In this case, the kinematic homology between the two elements allows for a straightforward mapping definition. Nonetheless, the implementation of intuitive BoMIs for nonhomologous architectures comes as a greater challenge.

Three major characteristics of human motion are of interest in the implementation of an intuitive body-machine mapping. Firstly, human motion involves several degrees of freedom (DOFs) and is achieved through redundant kinematics [10], [11], [5], often requiring preliminary feature selection and dimensionality reduction methods to extract meaningful information. A common approach consists in the adoption of *motion synergies*: functional sensory-motor modules exploiting the combination of several motion units [12]. Popular tools for the identification of motor synergies are signal compression algorithms such as principal component analysis (PCA) and linear discriminant analysis [13], [14]. Secondly, even common human actions like gait and planar reaching arm movements

Manuscript received: May 31, 2019; Revised: September 9, 2019; Accepted October 22, 2019.

This paper was recommended for publication by Editor Allison Okamura upon evaluation of the Associate Editor and Reviewers' comments.

This work was supported by the Swiss National Science Foundation (SNSF) and the European project RoboCom++.

The authors are with the Laboratory of Intelligent Systems (LIS), École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland. <https://lis.epfl.ch>

Corresponding author: matteo.macchini@epfl.ch

© 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

Digital Object Identifier (DOI): see top of this page.

present significant differences among individuals [15], [16], and cannot be easily generalized. Therefore, a personalized mapping could improve user's ability in teleoperation. Finally, movement is characterized by state-dependent nonlinearities in space, [17], [18], thus linear methods might not be sufficient to explain its complexity.

Among mobile robots, the aerial robotics field is undergoing a significant period of growth in industrial, security and research applications [19], [20]. Motion control paradigms have been proposed for the control of mobile robots using devices based on stereo vision for tracking motion of the operator's full body [21], [22] or portions of it such as face and hands [23], [24], [25]. A common feature of these interfaces is the fact that they consist of a predefined, discrete set of control commands, such as "take off", "go left". This approach to telerobotics is known as *supervisory control* [26]. Despite alleviating the user from the task of finely regulating each DOF of the robot, supervisory control does not allow fine and accurate maneuvering of the system. Moreover, the command set is normally defined in an arbitrary fashion or considering correspondences between human and robotic morphologies [27], thus not taking into account the specific operator's preferences. Data-driven body-machine interfaces have gained increasing attention over the last decade in a multitude of fields. Recent works investigated the natural interaction strategy of humans controlling drones through behavioral experiments [28], [29], to determine common mapping strategies, yet with a supervisory control paradigm. User-specific mapping functions have been proposed for anthropomorphic manipulators [16], [30] and ground robots [31]. Nonetheless, the first case is limited to homologous systems, while the second does not investigate the effect of such a method on learning and usability.

In a previous study, we focused on the implementation of an intuitive body-machine mapping for a fixed-wing UAV [6]. A study on human subjects led to the identification of two instinctive interaction behaviors: one consisting in torso movements, and a second one including arm motion. The two emerging behaviors were then tested in both simulation and hardware teleoperation, showing increasing proficiency compared to a remote controller. For the purpose of the project, a dedicated wearable hardware interface, the 'FlyJacket' [32], was designed. The previously implemented mapping strategy is derived from a limited set of users, and individual preferences are not taken into consideration. Despite the definition of a general method is desirable for several applications, it represents a suboptimal approach as it requires each person to adapt to the predefined interface.

In this article, we aim at reducing this effort by minimizing the difference between the mapping and the instinctive interaction behaviors of each individual, thus allowing a faster skill acquisition. We propose a novel framework capable of autonomously generating a user-adapted body-machine mapping function for robot operation (see Fig. 1). We apply machine learning methods to identify meaningful motion synergies from body movement and translate them into optimal commands for a drone. To compensate for motion errors, we investigated both linear and nonlinear methods. We assessed the framework capabilities through a pilot study,

showing that personalization and nonlinearity are features able to improve mapping accuracy. Subsequently, we generated personalized mappings for a set of participants and allowed them to teleoperate a virtual fixed-wing drone through a 3D path by using this custom interface. We additionally analyzed their motion to provide insight over humans' instinctive teleoperation strategies. Finally, we conducted a test on the teleoperation of a real drone to confirm the transferability of the generated mapping to a real-world scenario.

II. MAPPING LEARNING FRAMEWORK

In this section, we present the proposed framework and its application to the mapping definition between a human operator and a fixed-wing drone. The process is structured into four main steps, outlined in Fig. 2. First, kinematic data from a subject and a predefined trajectory from a simulated drone are acquired, while the user imitates the robot flight with arbitrary body motion. In a second step, relevant motion features are selected based on their correlation with robot behavior. Then, the features subset is further compressed into motion synergies, one per each available robot command. Finally, the identified synergies are transformed in commands for the drone through supervised regression.

Data acquisition and preprocessing: the user wears a head-mounted display (HMD) where a continuous flight trajectory is visualized in first-person view (FPV), corresponding to the frontal view of the simulated drone. The flight of the vehicle consists of eight subsequent segments: two roll (ϕ) maneuvers (lateral turns, right and left), two pitch (θ) maneuvers (vertical turns, up and down) and four mixed ones, where pitch and roll vary at the same time. ϕ and θ commands present the same shape for each maneuver, following: $A(1 - (\cos(t2\pi f)))(\pi/2)$ with $A = 0.5$, $T = 8$ s and $f = 1/T$. This particular list of maneuvers was chosen to cover the widest possible range of inputs that might be provided during the teleoperation of such robot. To make the trajectory easier to predict and follow, a red beam is visualized to show the future position of the drone in a 0.6 s time window, as shown in Fig. 3 (left). During this phase, the user has no control over the robot trajectory, and they are asked to move in an instinctive way, as if they were controlling the flight direction with their body. This approach is similar to the ones seen in [30], and [31]. During this phase, the human pilot is sitting on a stool. Indeed, subjects did not show a clear preference between a standing, sitting or lying position in our previous work [6]. Synchronously, we record the attitude of the UAV, expressed through roll (ϕ) and pitch (θ) values, at a sampling frequency of 300Hz. A schematic view of the scenario can be appreciated in Fig. 3. The participant's upper body is represented as a concatenation of nine different kinematic links interconnected by means of spherical joints: torso, shoulders, arms, forearms and hands, arranged in the kinematic chain displayed in Fig. 3 (left). We record 27 kinematic variables, corresponding to the Euler angles describing the orientation of each considered rigid body. We decided to track joint angles for simplicity, as they are in principle sufficient to reconstruct the limbs' positions, and to allow portability to wearable sensors such as inertial motion

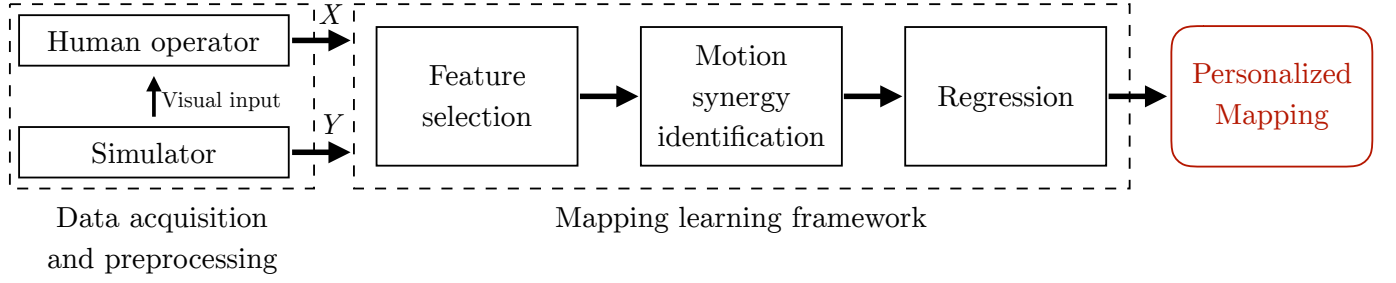


Fig. 2: Schematic overview of the framework. After the acquisition of the preset trajectory from the simulator (Y) and human movements (X), the most informative features are identified through a feature selection step, and further compressed into meaningful motion synergies. Supervised regression is employed to define a personalized mapping for the user.

units (IMUs), notably better suited for tracking angular than linear quantities. For the acquisition, the user is provided with a set of 25 reflective markers installed on a wearable velcro vest and tracked by a motion capture system. We denote with $\mathbf{x} = \{x_1, \dots, x_N\} \in R^N$ the set of kinematic variables acquired at each time step, and with $\mathbf{y} = \{y_1, \dots, y_M\}$ the robot attitude¹. Moreover, we denote with $\mathbf{X} = X_1, \dots, X_N$ and $\mathbf{Y} = Y_1, \dots, Y_M$ the set of time series respectively associated to \mathbf{x}, \mathbf{y} . All time series are normalized to have zero mean and unit variance. We perform one acquisition for each subject.

The human subject's visual perception delay during the observation of UAV trajectories could lead to time shifting between the kinematic variables and the robot attitude. Since our goal is to correlate these two datasets, time-shifting needs to be removed, as it deteriorates correlation. A first option would be to compensate for an average delay, corresponding to $\simeq 200\text{--}300\text{ ms}$ [33], which however would not be efficient in the case of intra-subject variability. Instead, we compress the kinematic time series \mathbf{X} and the robot actions \mathbf{Y} into the first component of a PCA, use it to compute the time shift by matching the maximum value and the rise time and apply this shift to all \mathbf{X} .

Feature selection: Since we make no assumptions on the user's instinctive motion, an automatic feature selection functionality has been implemented in order to extract the most relevant kinematic variables. For this purpose, we compute a quality metric for each kinematic variable x_i . The metric is given by two elements: the cross-correlation of each kinematic time series with the robot's roll and pitch and its signal-to-noise ratio (SNR). Firstly, we want to choose the kinematic variables from \mathbf{x} which correlate as much as possible with the given attitude \mathbf{y} . We compute the correlation of x_i with each robot command y_j through the cross-correlation coefficient K_{X_i, Y_j} :

$$K_{X_i, Y_j} = \frac{\sum_m (X_i[m] - \mu_{X_i})(Y_j[m] - \mu_{Y_j})}{\sigma_{X_i} \sigma_{Y_j}} \quad (1)$$

where $X_i[m]$ is the m -th sample of X_i . Note that K_{X_i, Y_j} is computed for a time $t = 0$ because we previously compensated for eventual time shifts. Also, $\mu_{X_i} = \mu_{Y_j} = 0$ and $\sigma_{X_i} =$

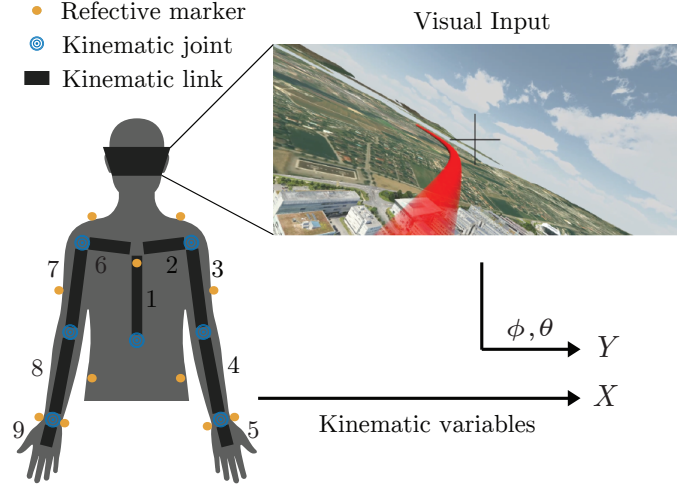


Fig. 3: Data acquisition scenario. The drone trajectory is given to the user as a visual input through the HMD, and they are free to move arbitrarily to follow the trajectory. The user's movements (X) and the roll and pitch values for the simulated drone (Y) are acquired synchronously. During this phase, the user's motion has no effect on the robot's trajectory.

$\sigma_{Y_j} = 1$ by normalization, yielding:

$$K_{X_i, Y_j} = \sum_m X_i[m] Y_j[m] \quad (2)$$

equivalent to the dot product “ \cdot ”. The total cross-correlation value for x_i is computed by summing its absolute correlation coefficient over all outputs:

$$K_{X_i, Y} = \sum_{j=1}^M |X_i \cdot Y_j| \quad (3)$$

Secondly, in order to take into account noise in the quality metric, we assume that X_i can be expressed as $X_i = S_i + N_i$ where N_i represents the noise, assumed to be fully caused by quantization, thus uniform and S_i is computed with a moving average Low-Pass FIR filter computed on 50 samples on X_i . Therefore, we denote with SNR_i the signal-to-noise ratio of the time series X_i . Then, the desired quality metric λ_i associated to a time series X_i can be expressed as:

$$\lambda_i = K_{X_i, Y} \cdot SNR_i \quad (4)$$

In summary, the quality metric λ_i is proportional to the amount of information relevant for the mapping definition which is

¹ $N = 27$: 3 Euler angles for 9 rigid bodies
 $M = 2$: roll and pitch values

carried by each kinematic variable x_i . The selection of the most informative kinematic variables is based on λ_i and done by conducting a Z-value outlier analysis with a threshold of 0.5. Since our goal is to define a function transforming the kinematic variables into drone commands of pitch and roll, we need to select at least M variables. Thus, if the feature selection produces less than M kinematic variables the M with the highest Z-values are selected. Therefore, let x^* denote the selected kinematic variables which are the most correlated to the drone's attitude and the ones mainly employed by the user. Thus, x^* represents the best subset to consider to transform into drone commands.

Motion synergy identification: in this step, both to simplify regression and to extract significant data, we run an additional data compression step to extract the user's motion synergies. Since PCA considers the correlation between the features and does not account for the robot behaviour we want to associate them with, we implemented a Canonical Components Analysis (CCA)[34]. The CCA provides M projections x_{SYN}^* of the selected features x^* , representing the functional motion synergies associated to roll and pitch commands.

Regression: finally, the identified motion synergies x_{SYN}^* , and the robot commands y are fed to a regression algorithm in order to define a mapping between the two, in a standard supervised learning problem. We investigated both linear and nonlinear solutions in order to compensate for the systematic motion errors made by humans. For the nonlinear mapping, we used Support Vector Regression (SVR) [35], implemented using the SKLearn Python library [36]. Hyperparameter optimization was run on each subject independently through grid search, with bounds limiting the estimator's variance, to allow a good interpolation between the sporadic data provided to the regression algorithm. The optimization is run through 10-fold cross-validation on the training dataset, down-sampled by a factor 10 to prevent overfitting.

III. RESULTS

In this section, the experimental protocols and results obtained through the study are summarized². First, we present a preliminary study conducted on pilot subjects used to validate the proposed architecture and quantify the effects of personalization and nonlinear methods in term of regression error. After, we describe the experiments conducted following the protocol summarized in Fig. 5. A first phase consisted in generating user-adapted mappings for a set of 16 participants, using the method outlined in Sect. II. Here we provide a study on the recognized motion patterns and synergies. In a second phase, we asked our subjects to control a simulated fixed-wing UAV using their personalized mapping and compared with state-of-art interfaces: a remote controller and a non-personalized BoMI. Finally, we used the same mappings to conduct hardware experiments with a subset of the subjects controlling a real quadrotor mimicking the dynamics of a fixed-wing drone.

²The experiments were approved by the École Polytechnique Fédérale de Lausanne Human Research Ethics Committee.

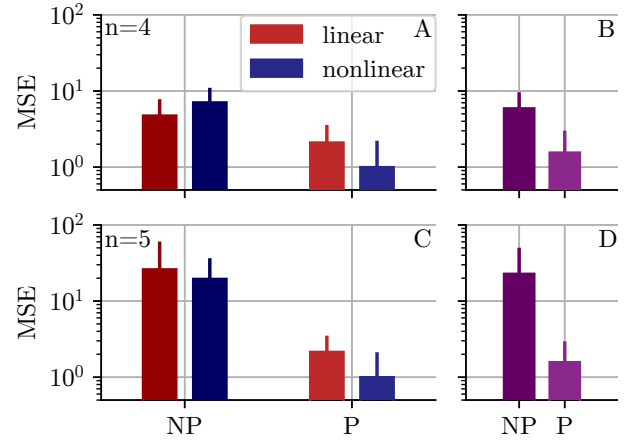


Fig. 4: Normalized regression error from pilot subjects. (B) and (D) show average values in (A) and (C) respectively. P and NP represent, respectively, personalized and non-personalized methods. In (A,B) only the 4 pilot subjects moving torso are considered. In (C,D), we include the outlier moving their right hand. Personalization shows clear advantage in terms of error, more importantly when all subject are included in the analysis. Nonlinear regressor shows similar effects in both cases, further improving the results.

Preliminary pilot study: we analyzed data collected on five pilot subjects to preliminarily assess the importance of personalization and nonlinearity on the definition of body-robot mapping. The two considered regression algorithms described in Sect. II were compared in terms of regression error. This metric allows estimating a nonlinearity score for motion data (the more difference between the two methods on a single dataset, the more nonlinear the observed system). Additionally, the effect of personalization was observed. We extrapolated a personalized mapping on each subject and compared its performance on the same subject and on all the remaining ones, in a leave-one-out logic. In Fig. 4, the regression error for both linear and nonlinear, personalized (P) and non-personalized (NP) approaches is shown. The minimal error is normalized to unit value. We made no prior assumptions on the subject's instinctive behavior. The framework aims to guarantee versatility for each individual, preventing the need for additional data analysis or parameter hand-tuning. For instance, a person may move the torso according to the drone's attitude, observed in [6] as a dominant behavior, while another may intuitively prefer to do it with one, or both hands. Both patterns should be equivalently decoded. During this phase of the experiment, four out of five subjects mimicked the drone's attitude using the torso, while one of them used only their right hand. For this reason, the figure is split into two parts: in Fig. 4 (A,B) only the four subjects moving the torso were included in the analysis, and the remaining one is considered an outlier. Contrarily, in Fig. 4 (C,D) the totality of the population is included. We see that, if only subjects using torso motion are taken into account, regressing on each individual lowers the average error by 78% as shown in Fig. 4 (B). Instead, if the subject who employed hand motion is included, the error drops by 96%, Fig. 4 (D). This difference is expected: including all subjects, we test a personalized mapping based on hand motion on subjects who did not focus on the movements of their

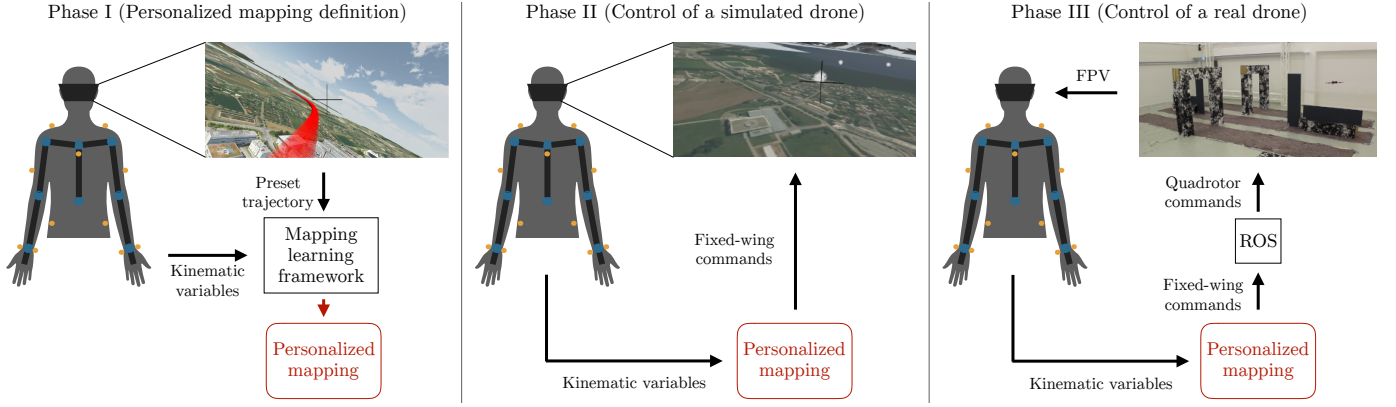


Fig. 5: Experimental design adopted for the validation of the proposed method: during Phase I, the framework is trained on the subject's instinctive motion strategy and translates it into robot controls. This procedure provides a mapping personalized on the user. Subsequently, teleoperation tests are performed in order to assess the learning curve of the given interface in simulation (Phase II) and a real quadrotor (Phase III).

right hand, resulting in big errors. In both cases, a nonlinear approach further reduced the error by more than 3 times, Fig. 4 (A,C).

Phase I - personalized mapping definition: 20 participants were recruited to verify the proposed algorithm. Each participant took part in the data acquisition procedure described in Sect. II, and our framework defined personalized mappings for each of them. First, we analyzed the quality metric. In Fig. 6 (A,B), we show the normalized cross-correlation $K_{X_i,Y}$ and $SNR_{i,Y}$, defined in Sect. II for each considered kinematic variable over all the subjects, in percentual value. Data show a statistical preponderance in the use of torso, as suggested in [6]. In total, 71.33% of the total cross-correlation, and 73.16% of the total SNR score is associated with torso angles. Fig. 6 (C) summarizes the presence of each kinematic variable in the identified selected subset x^* , as a result of the feature selection procedure for the totality of the experiments. For all participants (20, 100%) torso pitch (frontal flexion) was selected for the mapping. A large subset employed torso roll (lateral flexion), and a minority used torso yaw (rotation). Forearm motion was selected only for two subjects. For each participant, our framework was able to provide a mapping that was later used for teleoperation. This results proves the versatility of the proposed method for subjects moving in different fashions.

Phase II - control of a simulated drone: we run between-groups experiments to validate our method. 40 participants performed a teleoperation task in the simulator described in Sect. II, flying the fixed-wing UAV model through a 3D path consisting of 42 subsequent waypoints. The participants took part in this experiment directly after Phase I. The subjects were split into four groups, each using a different HRI:

- Group 1: a standard remote controller with a *non-personalized mapping*.
- Group 2: a BoMI implementing our previous *non-personalized mapping* described in [6].
- Group 3: a BoMI implementing the linear version of the proposed personalized method.
- Group 4: a BoMI implementing the nonlinear version of

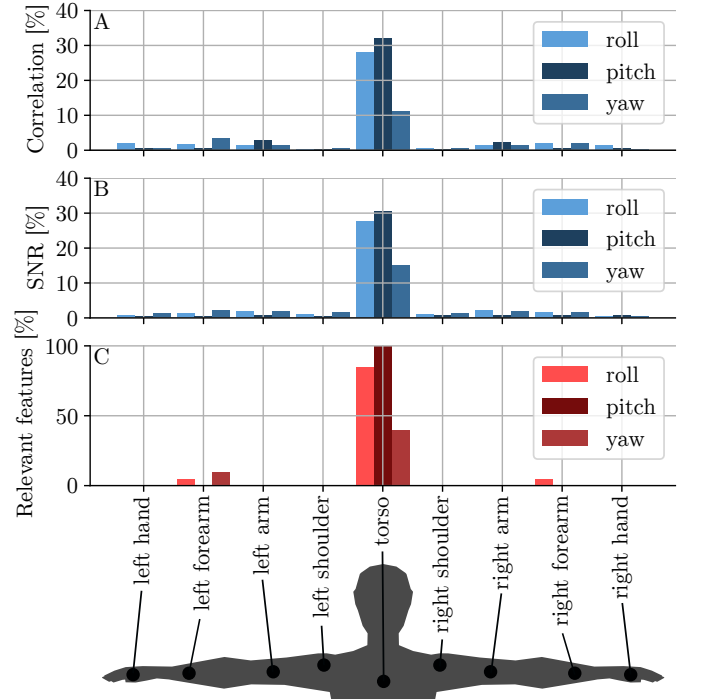


Fig. 6: The two components (A,B) of the quality metric for each feature show a statistical preponderance in the use of torso motion as a preferred motion behavior. (C) shows the features selected through the quality metric (4). Torso roll and pitch angles are chosen as relevant for most subjects, while torso yaw and forearms are selected only for a subset.

the proposed personalized method.

Note that Groups 3 and 4 correspond to the subjects mentioned in Phase I. Fig. 3 shows a schematic of the validation phases regarding both the simulated fixed-wing UAV and the real quadrotor UAV. All subjects were shown the simulation video and were let free to use their interface in an arbitrary way, as if they were controlling the flight of the fixed-wing UAV, but having no control over the drone's trajectory. No further instructions were provided. This phase is the same for all subjects to ensure coherent experimental conditions. Participants belonging to Groups 3 and 4 used their

personalized mapping interface, provided as an output of Phase I (see Fig. 5), to control the UAV in simulation. Afterwards, between Phase I and II, the participants belonging to Groups 1 and 2 were instructed how to use their interface, while the ones in Groups 3 and 4 were only asked to move coherently to Phase I.

All the 32 subjects were able to teleoperate the UAV in simulation through the waypoints. Each waypoint, visualized as a round cloud in the air (Fig. 5), is located at a fixed distance from the previous one, and a vertical or lateral displacement corresponding to an up, down, right or left turn. The turns are alternated in pseudo-random fashion in order for all subjects to perform the same amount of each. An accuracy score was computed to assess their ability. The distance from the n -th waypoint center, computed on the plane passing through it and perpendicular to the line crossing the $(n-1)$ -th and the $(n+1)$ -th waypoint, is used as a measure of the error. The distance is computed at the moment in which the UAV crosses the aforementioned plane.

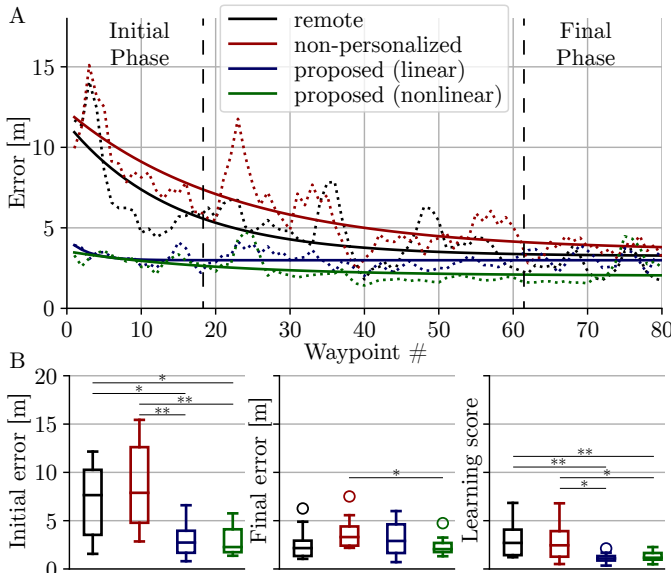


Fig. 7: (A) Error evolution during the teleoperation of the simulated drone. Dotted lines indicate the mean score across subjects in each group, continuous lines correspond to an exponential fit. Vertical dashed lines delimit respectively the end of initial phase and the start of the final phase. (B) Within-group initial error (left), final error (center) and learning score (right). Subjects provided with a personalized mapping have a better initial performance and less learning is required to them to reach the final performance plateau ($*p < 0.05$, $**p < 0.01$).

We let each participant teleoperate the simulated robot for two runs over the 42 waypoints, corresponding to a time of about five minutes, until they reached a performance plateau. In Fig. 7 (A), the evolution of the error throughout the two runs is shown as an estimate of the HRI learning curve. After the end of the second run, corresponding to waypoint 84, performance is stable for all groups. The first minute is where the most difference between groups happens, and it corresponds to the phase of maximal learning (i.e., maximal variation in the learning curve). We consider the error in this phase, considering the average number of waypoints

crossed during the first minute of teleoperation. For Groups 1 and 2, the error in this phase corresponds to $7.23 \pm 3.71m$ and $8.97 \pm 4.57m$, respectively. For Group 3, the error was reduced ($3.17 \pm 1.90m$, $p_{1,3}=0.019$, $p_{2,3}<0.01$), as for Group 4 ($2.98 \pm 1.56m$, $p_{1,4}=0.012$, $p_{2,4}<0.01$), proving the effectiveness of the proposed personalized method in the initial learning phase. Groups 3 and 4 show no significant difference, contrarily to our initial hypotheses and the pilot study ($p=0.88$). Data about this phase are shown in Fig. 7 (B, left). Likewise, Fig. 7 (B, center) shows performance in the final phase, defined as the last minute of teleoperation. No significant difference is observed among methods, excluding the pair non-personalized - proposed (nonlinear). We also evaluated the learning effects throughout the whole experiment. In Fig. 7 (B, right) we display the learning score, defined as the ratio between the performance error in the first and the last minute of the test. This score represents how much the subjects had to improve their performance to reach the final plateau. While this score is higher for Groups 1 ($3.32 \pm 2.15m$) and 2 ($2.87 \pm 1.93m$), it remains lower in Groups 3 and 4 ($1.18 \pm 0.52m$, $p_{1,3}<0.01$, $p_{2,3}=0.028$ and $1.29 \pm 0.53m$, $p_{1,4}<0.01$, $p_{2,4}=0.049$, respectively). Results show that the subjects using a personalized interface improved less (i.e., had to learn less) with respect to the ones using a remote or a wearable non-personalized ones. Statistical significance is tested with by Kruskal-Wallis test, due to the limited size of the participants pool.

After the experiment, we asked each subject personal feedback about the flight experience. In particular, we collected information about how tired (Q1) and disturbed by motion sickness (Q2) they felt at the end of the run, in a scale from 1 to 5. Data are shown in table I, and suggest that the participants using the BoMI, in average, did not feel significantly more tired and experienced less motion sickness than participants using a remote controller.

	Group 1	Group 2	Group 3	Group 4
Q1	2.00 ± 1.07	2.16 ± 1.07	2.0 ± 0.82	1.33 ± 0.66
Q2	2.86 ± 0.99	2.33 ± 1.49	1.83 ± 1.21	1.67 ± 1.25

TABLE I: Subjective feedback from the participants on experienced fatigue (Q1) and motion sickness (Q2) during Phase II in the form [mean \pm std], scale 1 to 5.

Phase III - control of a real drone: as a final step, we assessed the ability of transferring the teleoperation skills acquired during the simulation test to hardware implementation. This phase follows the same protocol as in the real-world experiments performed in [6]. 5 subjects from Group 1 and 5 subjects from Group 3 controlled the flight of a quadrotor through a path consisting in four obstacles for a total of three runs. The test was performed one week later than the Phase II study. The quadrotor controls were set to mimic fixed-wing dynamics and fly at constant speed of $0.3m/s$ in a controlled environment. The user was allowed, through the same mapping provided during the previous phase, to issue roll and pitch commands, translated respectively into vertical velocity and a combination of roll and yaw rate in the quadrotor body frame.

The experiments were performed with a Parrot Bebop 2.0

quadrotor connected through ROS to a ground control station. The subjects were provided with the same motion capture markerset as described previously. Their motion, converted into roll and pitch controls through their personalized mapping, was transmitted at a frequency of $10Hz$ to the drone. A safety fence was installed in the tracked volume to separate the zone dedicated to the user from the flight area.

All the subjects from Group 3 were able to control the robot's trajectory (Fig. 8) through the path with a total of $92.5 \pm 6.13\%$ obstacle avoidance success, confirming the transferability of the learned teleoperation skills in a real-world environment.

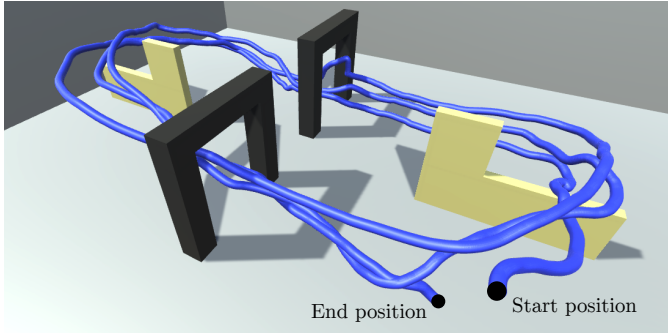


Fig. 8: 3D reconstruction of the hardware test scenario. In blue, the trajectory of the real drone for a sample subject.

IV. DISCUSSION

In this paper, we describe the implementation and validation of a framework able to learn relevant motion patterns and mapping functions for a body-machine interface. We applied it to the case of teleoperation of a fixed-wing drone. Moreover, through preliminary tests, we showed that the proposed method can analogously be applied to the control of a real quadcopter. A quantitative analysis of subjects' performance regarding the teleoperation of a real drone is out of the scope of this letter and it will be addressed in future work. The mapping results thus optimized for the specific user, taking into account their instinctive motion when asked to behave as if they were controlling a drone's motion in FPV. We demonstrate by preliminary experiments that personalization and nonlinearity are effective features for the implementation of an improved BoMI, as in Fig. 4. Based on this, both linear and nonlinear regression methods have been validated. Through the extraction of functional motion synergies specific to the task, the proposed architecture is able to provide insights on humans' preferred motor actions for teleoperation. Specifically, the torso has been identified by our framework as the dominant body segment instinctively adopted by users for robot control (Fig. 6), confirming the results shown in a previous study [6].

In addition, we observed that participants using a standard remote controller and a non-personalized BoMI took several minutes to reach stable performance. Instead, naive users provided with a personalized body-machine mapping showed no significant learning patterns throughout the whole experiment, reaching a comparable proficiency plateau from the start. In the

first phase of the test, where a significant portion of learning takes place, the subjects using a personalized body-machine mapping performed 57% better than the ones using a remote controller and 65% better than non-personalized BoMI users (see Fig. 7). These results suggest that the personalization of the interface mitigated for the initial learning effort, effect typically achieved with the adoption of augmented feedback to the user.

However, participants provided with linear and nonlinear mapping functions achieved very similar performance. It is important to remark that our preliminary study (Fig. 4) provides an estimation of the motion nonlinearity through regression error assessment, but our performance measure is not directly related to this error: small nonlinearities might still be easy to correct for a user. Indeed, most of the previous studies in the field [16], [17], [18] focus on reaching arm movements, while experimental subjects employed mainly torso motion. Being the torso a larger body segment than arms, and belonging to a less peripheral area, it could be able to provide stronger proprioceptive and vestibular feedback, and thus be easier to control in a more linear manner. Moreover, torso movements directly affect head motion, where most of the human sensory organs are situated. Additionally, the limited amount of controllable degrees of freedom of the robot allowed all users to reach low error levels, especially with personalization and training. We can interpret this as a 'saturation' effect, which cannot be further improved by acting solely on the interface.

Our experiments were conducted in a motion capture system to track the participant's body. However, our method is not limited to be used in such environments. Conversely, the recognized motion synergies and relevant body segments can help its adaptation for outdoor implementation, for instance, by means of IMUs as in [32]. They can also be used for optimal sensor placement. Such application could facilitate the deployment of body-controlled robotic systems on the field, providing at the same time a minimal volume to be transported for the wearable instrumentation.

The results of this work open several interesting future research directions. It is important to remark that the framework's mathematical formulation makes no specific assumptions on the nature of either the source of the commands (here, upper body motion) nor the morphology of the controlled robotic system (here, a fixed-wing UAV). Extending to a different body configuration, for instance adapted for impaired individuals, or a different input device, or to more complex robotic morphologies is an exciting future research direction to be investigated. Additionally, investigating online learning techniques for real-time adaptation of the mapping could be a viable option to cope with a user's change in preferences, or with their improvements with experience. Moreover, although our method was able to provide a subject-adapted mapping for all the observed motion patterns, we individuated a small diversity in the users' behaviors. Therefore, testing our approach on a larger population will undoubtedly add value to our work.

V. CONCLUSIONS

This study represents a significant contribution in the domain of human-robot interfaces applied to teleoperation. The use of remote controllers is an established method to control mobile robots. Nonetheless, They provide a limited capability of adaptation, requiring a significant amount of training time and cognitive effort to be proficiently mastered. The exploitation of natural body motion and the implementation of an intelligent framework capable of translating intuitive movements into robot commands could make the control of complex robotic systems a user-friendly experience also for novices. Furthermore, the generality of this method could allow to extend it to a broader population and different robot morphologies.

REFERENCES

- [1] G. Niemeyer, C. Preusche, S. Stramigioli, and D. Lee, "Telerobotics," in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Springer, Berlin, Heidelberg, 2016, pp. 1085–1108.
- [2] K. Nagatani, S. Kiribayashi, Y. Okada, K. Otake, K. Yoshida, S. Tadokoro, T. Nishimura, T. Yoshida, E. Koyanagi, M. Fukushima, and S. Kawatsuma, "blackEmergency response to the nuclear accident at the Fukushima Daiichi Nuclear Power Plants using mobile rescue robots," *J. Field Robot.*, vol. 30, no. 1, pp. 44–63, 2013.
- [3] R. R. Murphy, S. Tadokoro, D. Nardi, A. Jacoff, P. Fiorini, H. Choset, and A. M. Erkmen, "Search and rescue robotics," in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Springer, Berlin, Heidelberg, 2008, pp. 1199–1222.
- [4] R. H. Taylor, A. Menciassi, G. Fichtinger, P. Fiorini, and P. Dario, "blackMedical Robotics and Computer-Integrated Surgery," in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Springer, Berlin, Heidelberg, 2008, pp. 1199–1222.
- [5] M. Casadio, R. Ranganathan, and F. A. Mussa-Ivaldi, "blackThe Body-Machine Interface: A New Perspective on an Old Theme," *J. Mot. Behav.*, vol. 44, no. 6, pp. 419–433, 2012.
- [6] J. Miehlebradt, A. Cherpillod, S. Mintchev, M. Coscia, F. Artoni, D. Floreano, and S. Micera, "blackData-driven body-machine interface for the accurate control of drones," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 115, no. 31, pp. 7913–7918, 2018.
- [7] S. Kim, C. Kim, B. You, and S. Oh, "blackStable whole-body motion generation for humanoid robots to imitate human motions," in *IEEE Intl. Conf. on Intell. Rob. and Sys. (IROS)*, Oct 2009, pp. 2518–2524.
- [8] B. Brggemann, T. Röhling, and J. Welle, "blackCoupled Human-machine Tele-manipulation," *Procedia Manuf.*, vol. 3, pp. 998–1005, 2015.
- [9] L. Sentis and O. Khatib, "blackA whole-body control framework for humanoids operating in human environments," in *IEEE Intl. Conf. on Rob. and Aut. (ICRA)*, 2006, pp. 2641–2648.
- [10] B. M. Cruse H, "blackThe human arm as a redundant manipulator: the control of path and joint angles," *Biol. Cybern.*, vol. 57, no. 1-2, pp. 137–144, 1987.
- [11] O. Khatib, E. Demircan, V. D. Sapio, L. Sentis, T. Besier, and S. Delp, "blackRobotics-based synthesis of human motion," *J. Physiol. Paris*, vol. 103, no. 3, pp. 211 – 219, 2009.
- [12] A. Leo, G. Handjaras, M. Bianchi, H. Marino, M. Gabiccini, A. Guidi, E. P. Scilingo, P. Pietrini, A. Bicchi, M. Santello, and E. Ricciardi, "blackA synergy-based hand control is encoded in human motor cortical areas," *eLife*, vol. 5, p. e13420, 2016.
- [13] R. Vinjamuri, V. Patel, M. Powell, Z.-H. Mao, and N. Crone, "blackCandidates for Synergies: Linear Discriminants Versus Principal Components," *Comput. Intell. Neurosci.*, vol. 2014, pp. 9–19, 2014.
- [14] T. D. Sanger, "blackHuman Arm Movements Described by a Low-Dimensional Superposition of Principal Components," *J. Neurosci.*, vol. 20, no. 3, pp. 1066–1072, 2000.
- [15] N. Stergiou and L. M. Decker, "blackHuman movement variability, nonlinear dynamics, and pathology: Is there a connection?" *Hum. Mov. Sci.*, vol. 30, no. 5, pp. 869 – 888, 2011.
- [16] R. M. Pierce and K. J. Kuchenbecker, "blackA data-driven method for determining natural human-robot motion mappings in teleoperation," in *IEEE Intl. Conf. on Biomed. Robot. and Biomech. (BioRob)*, June 2012, pp. 169–176.
- [17] R. Shadmehr and F. Mussa-Ivaldi, "blackAdaptive representation of dynamics during learning of a motor task," *J. Neurosci.*, vol. 14, no. 5, pp. 3208–3224, 1994.
- [18] M. F. Ghilardi, J. Gordon, and C. Ghez, "blackLearning a visuomotor transformation in a local area of work space produces directional biases in other areas," *J. Neurophysiol.*, vol. 73, no. 6, pp. 2535–2539, 1995.
- [19] D. Floreano and R. J. Wood, "blackScience, technology and the future of small autonomous drones," *Nature*, vol. 521, p. 460466, 2015.
- [20] SESAR. European drones outlook study.
- [21] K. Ikeuchi, T. Otsuka, A. Yoshii, M. Sakamoto, and T. Nakajima, "blackKinectDrone: Enhancing Somatic Sensation to Fly in the Sky with Kinect and AR.Drone," in *ACM Augm. Hum. Intl. Conf. (AH)*, 2014, pp. 53:1–53:2.
- [22] A. Sanna, F. Lamberti, G. Paravati, and F. Manuri, "blackA Kinect-based natural interface for quadrotor control," *Entertain. Comput.*, vol. 4, no. 3, pp. 179 – 186, 2013.
- [23] A. Sarkar, K. A. Patel, R. K. G. Ram, and G. K. Kapoor, "blackGesture control of drone using a motion controller," in *IEEE Intl. Conf. on Indust. Inform. and Comput. Sci. (CIICS)*, 2016, pp. 1–5.
- [24] T. Naseer, J. Sturm, and D. Cremers, "FollowMe: Person following and gesture recognition with a quadcopter," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nov. 2013, pp. 624–630.
- [25] J. Nagi, A. Giusti, G. A. D. Caro, and L. M. Gambardella, "Human Control of UAVs using Face Pose Estimates and Hand Gestures," in *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Mar. 2014, pp. 1–2.
- [26] T. B. Sheridan, "blackAdaptive Automation, Level of Automation, Allocation Authority, Supervisory Control, and Adaptive Control: Distinctions and Modes of Adaptation," *IEEE Trans. Syst., Man, Cybern.*, vol. 41, no. 4, pp. 662–667, 2011.
- [27] H. J. Chang, T. Fischer, M. Petit, M. Zambelli, and Y. Demiris, "blackKinematic Structure Correspondences via Hypergraph Matching," in *IEEE Conf. Comput. Vis. Pattern. Recognit.*, 2016, pp. 4216–4225.
- [28] K. Pfeil, S. L. Koh, and J. LaViola, "blackExploring 3D Gesture Metaphors for Interaction with Unmanned Aerial Vehicles," in *ACM Intl. Conf. on Intell. User Interf. (IUI)*, 2013, pp. 257–266.
- [29] J. R. Cauchard, J. L. E. K. Y. Zhai, and J. A. Landay, "blackDrone & Me: An Exploration into Natural Human-drone Interaction," in *ACM Intl. Joint Conf. on Pervas. and Ubiqu. Comput. (UBICOMP)*, 2015, pp. 361–365.
- [30] R. P. Khurshid and K. J. Kuchenbecker, "blackData-Driven Motion Mappings Improve Transparency in Teleoperation," *Presence (Camb.)*, vol. 24, no. 2, pp. 132–154, 2015.
- [31] C. Melidis, H. Iizuka, and D. Marocco, "blackIntuitive control of mobile robots: an architecture for autonomous adaptive dynamic behaviour integration," *Cogn. Process.*, vol. 19, no. 2, pp. 245–264, 2018.
- [32] C. Rognon, S. Mintchev, F. I. T. Dell'Agnola, A. T. Cherpillod, D. Atienza Alonso, and D. Floreano, "blackFlyJacket: An Upper Body Soft Exoskeleton for Immersive Drone Control," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 2362–2369, 2018.
- [33] R. Mansfield, "blackLatency functions in human vision," *Vis. Res.*, vol. 13, no. 12, pp. 2219 – 2234, 1973.
- [34] W. K. Hardle and L. Simar, "blackCanonical Correlation Analysis," in *Applied Multivariate Statistical Analysis*. Springer, Berlin, Heidelberg, 2015.
- [35] H. Drucker, C. J. C. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, "blackSupport Vector Regression Machines," in *IEEE Adv. Neural Inf. Process. Syst.*, 1997, pp. 155–161.
- [36] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg et al., "blackScikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.