Feedback-Aided Data Acquisition Improves Myoelectric Control of a Prosthetic Hand

Andrea Gigli[1], Donato Brusamento[1], Roberto Meattini[2], Claudio Melchiorri[2], and Claudio Castellini[1]

[1]Institute of Robotics and Mechatronics, German Aerospace Center (DLR), Wessling, Germany
[2]Department of Electrical, Electronic and Information Engineering (DEI), University of Bologna, Bologna, Italy

E-mail: andrea.gigli@dlr.de

September, 2020

Abstract.

Pattern-recognition-based myocontrol can be unreliable, which may limit its use in the clinical practice and everyday activities. One cause for this is the poor generalization of the underlying machine learning models to untrained conditions. Acquiring the training data and building the model more interactively can reduce this problem. For example, the user could be encouraged to target the model's instabilities during the data acquisition supported by automatic feedback guidance. Interactivity is an emerging trend in myocontrol of upper-limb electric prostheses: the user should be actively involved throughout the training and usage of the device.

In this study, 18 non-disabled participants tested two novel feedback-aided acquisition protocols against a standard one that did not provide any guidance. All the protocols acquired data dynamically in multiple arm positions to counteract the limb position effect. During feedback-aided acquisition, an acoustic signal urged the participant to hover with the arm in specific regions of her peri-personal space, de facto acquiring more data where needed. The three protocols were compared on everyday manipulation tasks performed with a prosthetic hand. Our results showed that feedback-aided data acquisition outperformed the acquisition routine without guidance, both objectively and subjectively, indicating that interaction during the data acquisition is fundamental to improve myocontrol.

Keywords: myoelectric control, training data acquisition, feedback guidance, limb position effect, online machine learning, prosthetic hand

Submitted to: *J. Neural Eng.*

1. Introduction

The loss of an upper limb can affect the ability to carry out essential activities of daily living (ADLs) [1]. A variety of prosthetic devices are clinically available to restore the missing limb's functionalities and eventually improve the person's autonomy, health, and lifestyle [2]. According to recent surveys, however, between 10 % and 40 % of the people living with a limb-loss renounce using active prostheses [3, 4]. While these statistics may be related to the etiology and the level of amputation, commonly reported causes of prosthesis rejection include lacking functionalities or unintuitive and unreliable control.

Myocontrol approaches based on pattern recognition (PR) have emerged to address typical demands of prostheses' users, such as controlling multiple hand functions, regulating the grip strength, and overcoming the complex mode-switching mechanism of sequential myocontrol [5]. PR-based approaches allow the user to produce seamless transitions between multiple grasp patterns, or even to simultaneously and proportionally (s/p) control multiple degrees of freedom (DoFs) of the prosthesis [6, 7]. A shortcoming of PR models is that their reliability is compromised when used in conditions different from the training conditions. This is often the case during myoelectric devices daily use, due to the variability of the myoelectric signal and its measurements, e.g., surface electromyography (sEMG). Sources of variability include, for example, variations of the skin connectivity, electrodes placement, limb position, as well as fatigue phenomena, and the evolution of the user's cognitive capabilities [8].

An intuitive way to improve the robustness of the control model is to capture the variability of the sEMG signal in the training data and leverage the learning algorithm's generalization capabilities. For example, data acquisition can be designed to record myoelectric data for different limb orientations or electrodes positioning. Both batch and adaptive approaches can be used to collect the training data. Adaptive learning allows updating the model with new data upon necessity, reducing the effort to forecast all the possible sources of sEMG variability at once. Numerous studies confirm the benefits of improving myocontrol by progressively refining the training dataset [8, 9, 10] and a commercially available PR-based myocontrol system, Control Coach by COAPT†, allows for incremental model updates.

We believe that the reason for the success of this approach lies in its interactivity. The user is expected to evaluate the model performance and actively address its flaws by collecting appropriate training data. This can be seen as a special case of the human-in-the-loop paradigm [11, 12], in which bidirectional user-prosthesis interaction is enforced. Involving the user in the myocontrol loop both during the training and testing of the prosthesis enables faster understanding and embodiment of the device and favors the production of more effective control signals [13, 14].

The deployment of data acquisition routines outside the laboratory could benefit from providing automatic guidance to the user during the acquisition routine. In fact, without supervision from an expert, the user may be unable to identify and address the model's weaknesses, resulting in fruitless or even harmful model updates. For example, the user can be guided by precisely structuring the data collection routine or providing instantaneous feedback on the model's performance during data collection. To this aim, Woodward and Hargrove [15] designed a virtual reality game to structure a multi-arm-position data acquisition and provide instantaneous visual feedback of the myocontrol performance. They showed that guidance via serious games increases users autonomy throughout the model adaptation process. Hahne et al. [16] highlighted that the effectiveness of adaptive myocontrol could be enhanced by overcoming the typical separation between data acquisition and model updates. They proposed a data acquisition protocol in which the model was updated and evaluated online, while users received instantaneous feedback about the evolution of the model performance. This allowed users to identify and address the model's limitations already during the data acquisition and achieve robust myocontrol performance in a Fitt's law test with few model updates.

In this work, we focus on alleviating the negative effect of limb position variations on the myocontrol of a prosthetic hand. A detailed characterization of this problem can be found in [17]. Proposed solutions include designing sEMG features [18] or control models [19] that are less sensitive to the limb position effect, or acquiring training data in multiple arm positions. Multi-position acquisition can be performed statically, by repeating the target hand gestures in different arm configurations [20], or dynamically, by executing predefined arm movements [21, 22]. Multi-position training proved to enhance myocontrol performance compared to single-position training, and dynamic protocols also reduce the time and effort needed to complete the data acquisition [23]. None of the multi-position acquisition protocols found in the literature provides feedback guidance to the best of our knowledge. Prosthesis users are typically required to perform a movement routine without being fully aware to which extent each arm position contributes

†`https://coaptengineering.com/control-coach`

to improving the myocontrol model. We argue that identifying in realtime which arm configurations are most critical for the model and signaling them to the user would improve the data acquisition efficiency.

We propose a novel protocol to collect sEMG data dynamically in multiple arm positions under automatic feedback guidance. The protocol combines data acquisition, online model building, and instantaneous feedback about the usefulness of the recorded data. We compare two variants of the novel feedback-aided acquisition protocol to a standard one that does not provide feedback guidance. All the protocols build the model online and are based on the dynamic data acquisition described in [23]. Both feedback-aided protocols adopt the same feedback mechanism, but one of them also implements automatic sample selection to discard unnecessary training samples and reduce the number of model updates.

2. Materials and Methods

This study evaluates the effects of using a feedback signal to guide the acquisition of training data for myoelectric controllers of prosthetic hands. The performances of two feedback-aided data acquisition procedures and one standard acquisition using no feedback guidance were compared based on the controllability of a prosthetic hand during a series of realistic manipulation tasks.

2.1. Participants

Eighteen non-disabled persons (aged 26.3 ± 4.6 years, 16 men and 2 women) participated in the experiment. Twelve participants had no prior experience in myoelectric control, while six had already used myoelectric prosthetic hands in previous user studies. Every participant received an oral and written description of the experiment and signed an informed consent form. The study was conducted at the German Aerospace Center (DLR) according to the WMA Declaration of Helsinki and approved by DLR's internal committee for personal data protection.

2.2. Experimental Setup

The muscular activity of the forearm of the dominant arm was measured using a Myo armband‡ by Thalmic Labs placed about 5 cm below the elbow. The bracelet comprised eight sensors, each recording an sEMG signal at a sampling rate of 200 Hz. A standard quick-release prosthetic connector fixed to a wrist/hand orthotic splint made it possible to anchor the prosthesis at the extremity of sound limbs. An i-LIMB Ultra Revolution prosthetic hand§ by Touch Bionics (now Össur) allowed independent flexion/extension of the five fingers and abduction/adduction of the thumb through six motors under direct current control. The devices communicated via a serial-port-over-Bluetooth with a laptop used to run the myocontrol software. The acoustic feedback was reproduced using the speakers of the laptop. A custom software suite written in the C# language provided the graphical interface to coordinate the data acquisition, labeled and processed sEMG data, generated the feedback signal, and implemented realtime myocontrol.

The experiment took place in a domestic-like laboratory environment. We arranged several household objects on a table, two shelves, and on the floor. We placed the table 40 cm next to the shelves, and we regulated its height to match the waist level of each participant. The shelves were 40 cm and 150 cm high. The study was videotaped in order to measure the participant's performance after the experiment. Figure 1A shows the experimental setup.

2.3. Incremental model building

The sEMG readings were preprocessed in realtime upon collection. The measurement from each of the 8 channels was rectified, computing its absolute value, and low-pass filtered using a second-order Butterworth filter with a cutoff frequency of 1 Hz.

The data acquisition software labeled incoming training samples with the activation commands for the motors of the prosthetic hand's fingers. Each command consisted of a normalized velocity ranging between 0 and 1, corresponding to extending or flexing the finger with maximum speed. Since all the hand gestures considered in this experiment could be realized by controlling one subset of the fingers with the same velocity command, the model had effectively 3 DoFs.

Training data was collected only for extreme velocity command values, that is, for hand gestures in which each finger was either fully extended or fully flexed. Intermediate velocity commands were excluded because they could lead to inaccuracies in the recorded data due to the participants' different reaction times [24]. Previous works, such as [25, 23], showed that regression models resulting from this training procedure still yield effective s/p control.

To provide appropriate feedback guidance during the data acquisition, it was necessary to incorporate each training sample into the model quickly upon collection. Therefore, we trained the s/p control model using an instance of incremental ridge regression (iRR) with random Fourier features (RFFs). iRR builds a

‡`https://support.getmyo.com/hc/en-us/articles/203398347-Getting-started-with-your-Myo-armband`

§`https://www.ossur.com/en-us/prosthetics/arms/i-limb-ultra` regression model incrementally by computing rank-one model updates when new training data is available. The iRR formulation allowed us to update the model and generate predictions with bounded time and space complexity. RFFs is a nonlinear mapping of the input space into a high-dimensional feature space obtained by using sinusoidal basis functions that have randomly sampled frequencies. By drawing those frequencies from an adequate probability distribution and choosing a sufficiently high mapping dimensionality, iRR with RFF approximates ridge regression with a Gaussian kernel [25]. Consequently, RFFs extend the capacity of iRR to perform nonlinear regression while maintaining the properties of incrementality and boundedness of the model update. This is relevant in applications that require online learning of nonlinear regression models and has proven beneficial for s/p myocontrol applications [25, 26, 27, 23]. A detailed description of iRR-RFF can be found in [25]. The prediction function of iRR-RFF is

$$\hat{\mathbf{y}} = \mathbf{W} \cdot \Phi(\mathbf{x}) \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^d$ is an input sample, $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^D$ is a nonlinear RFF mapping, $\mathbf{W}$ is an $M \times D$ matrix of scalar weights, and $\hat{\mathbf{y}} \in \mathbb{R}^M$ is the computed prediction. In this experiment, training pairs $\{\mathbf{x}, \mathbf{y}\}$ consisted of an sEMG measurement and the corresponding velocity commands for the fingers' motors. The input and output dimensionality of the model were $d = 8$ and $M = 3$. The regularization parameter λ of the ridge regression was set to 1, while the bandwidth γ and the dimensionality D of the RFF mapping were set to 0.1 and 300, respectively. The model weights were initialized to zero before the data acquisition, $\mathbf{W} = \mathbf{0}_{M,D}$.

2.4. *Acoustic feedback and sample selection*

The idea behind feedback-aided data acquisition is to guide the participant with an appropriate feedback signal in order to maximize the amount of informative and non-redundant data collected in a fixed amount of time. In an online learning problem, the informativeness of a correctly-labeled training sample $\{\mathbf{x}, \mathbf{y}\}$ for the model can be evaluated based on the prediction error

$$e_p(\mathbf{y}, \hat{\mathbf{y}}) = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 \tag{2}$$

where $\hat{\mathbf{y}}$ is the prediction of the sample using the model. A small prediction error indicates that the model can accurately predict the label for that training sample and, therefore, the sample might be redundant for the model. A significant prediction error, instead, indicates that the model fails to predict the right label and may improve by integrating the sample. For the sake of clarity, we omit the argument of the prediction error in the remainder of the paper.

2.4.1. Feedback signal For our purposes, we designed an acoustic feedback signal with a fixed tone and variable volume. The volume of the signal ranged between 0 and a maximum value V and varied proportionally with the prediction error according to

$$f(e_p) = \max\left\{0, \min\left\{a e_p^2 + \frac{V - a\theta_u^2}{\theta_u} e_p, V\right\}\right\} \tag{3}$$

in which a was a scalar regulating the quadratic relation between error and volume, and θ_u was a threshold related to the prediction error. We set the values of the parameters to $V = 0.5$, $a = 70$, and $\theta_u = 0.05\sqrt{3}$. The value of θ_u corresponded to 5 % of the maximum theoretical value of the prediction error in our experiment, which was predicting an open hand gesture instead of a power grasp gesture.

2.4.2. Sample Selection We also used the prediction error to discard possibly redundant training samples. We defined a sample selection criterion to update the model only with those training samples for which

$$e_p \geq \theta_u \tag{4}$$

where θ_u is the update threshold defined before.

2.5. *Experimental protocol*

Every participant tested all the data acquisition strategies. We counterbalanced possible learning effects by administering the strategies to the participants in randomized order. We assigned each of the six permutations of the training conditions to one experienced and two naive participants picked at random. After each data acquisition, the resulting myocontrol model was tested in a sequence of realtime manipulation tasks. Participants repeated the sequence of tasks three times. The first two repetitions of the sequence allowed the participants to familiarize themselves with the prosthetic system and the myocontrol model, while the third one was used to measure the myocontrol performance. For this reason, we referred to the third repetition of the task sequence as a performance evaluation session.

2.5.1. Data acquisition All the data acquisition strategies required the participants to perform several target hand gestures while moving their arm in the reachable space. We selected three target hand gestures: namely a power grasp, a resting hand, and an index pointing. The selection was based on their relevance in ADLs, according to the literature [28]. The target hand gestures were acquired in the order reported above in every data acquisition. Since the myocontrol model was built incrementally, the

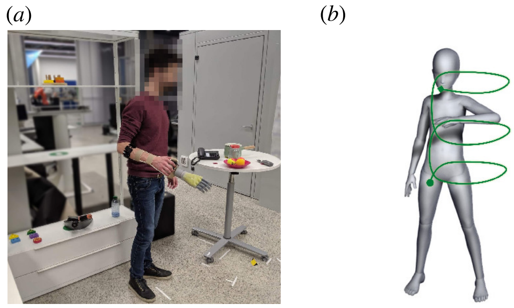


Figure 1. *Experimental setup and arm motion during data acquisition.* *(a)* The prosthetic system comprised a Myo armband by Thalmic Labs for sEMG reading, and an i-LIMB Ultra Revolution prosthetic hand by Össur. The experimental setup included common household objects placed onto one table and two shelves. The speakers of the control laptop provided acoustic feedback. *(b)* Participants wore the prosthetic system throughout the data acquisition. Every data acquisition routine required to perform several target hand gestures while moving the arm in a predefined trajectory. The motion proceeded from the circle to the square with the palm oriented downward and continued in the opposite direction with the palm oriented upward.

use of different orders would have possibly led to incomparable models.

Before the experiment, participants were explained the data acquisition protocols and were asked to practice them. Emphasis was put into enforcing a consistent arm movement across participants and strategies. Meanwhile, the volume of the speakers was regulated so to ensure that the feedback was distinctly audible. Nonetheless, the experimenter supervised the data acquisition and provided direct guidance when the participants performed the arm movement at the wrong pace or ignored the acoustic feedback.

Participants donned the prosthetic system on the dominant arm at the beginning of the experiment, and no adjustment of the sensors was allowed after that. Wearing the prosthesis during the data acquisition reduced the differences between the training and testing conditions caused by factors such as the electrodes' placement and the weight of the prosthetic device.

No-Feedback Data Acquisition (NF-DA) The data acquisition routine without feedback guidance adapted the dynamic acquisition presented in our previous work [23] to the setup of this study. Participants performed each target hand gesture while moving their arm in a predefined trajectory. During the procedure, they did not receive any feedback. The model was built online with each new training sample, as detailed in section 2.3. The trajectory uniformly covered the reachable space of the participant with a helical movement, Figure 1B. The movement was performed with constant speed from the level of the waist to the level of the head with the palm oriented downward; it continued in the opposite direction with the palm oriented upward. This whole sequence was repeated twice, without interruptions. The motion lasted 45 s for each hand gesture and took 135 s in total. The procedure is synthesized in Algorithm 1.

Feedback-Aided Data Acquisition (FA-DA) The data acquisition routine with feedback guidance extended NF-DA with the acoustic feedback detailed in section 2.4.1. The acquisition software used the incoming training samples to generate the acoustic feedback and to build the myocontrol model in realtime. Participants had to perform the desired grasp and follow the usual arm trajectory while modulating the arm's velocity based on the feedback. They should proceed with the same speed used during NF-DA when the feedback was not audible and hover with the arm in the areas where the feedback intensity increased. Since the feedback was proportional to the prediction error, this procedure led the participants to collect more data in critical arm configurations. The model incrementality prevented participants from slowing down indefinitely in critical areas of the reachable space. Training samples were continuously integrated into the myocontrol model, which immediately reduced the prediction error and, consequently, the volume of the feedback signal. The acquisition of each gesture lasted 45s. Differently from NF-DA, however, participants were not expected to cover the whole trajectory twice per gesture. The procedure is synthesized in Algorithm 2.

Feedback-Aided Data Acquisition with Sample Selection (FASS-DA) The data acquisition routine with feedback guidance and sample selection was obtained by integrating FA-DA with the sample selection criterion described in section 2.4.2. All the incoming training samples were used to generate the acoustic feedback, but only a limited number of non-redundant samples were selected and used to build the myocontrol model in realtime. Participants perceived no formal difference between the two feedback-aided acquisition routines. The procedure is synthesized in Algorithm 3.

2.5.2. Realistic myocontrol tasks After every data acquisition, the resulting myocontrol model was tested by engaging the participants in a series of five manipulation tasks. The tasks were inspired by realistic ADLs proposed in assessment protocols for prosthetic control, such as ACMC [29] and SHAP [30]. The tasks are described in Table 1. In the case of bimanual tasks, we assigned each action of the task either to the prosthetic hand or the sound hand.

Algorithm 1: No-Feedback Data Acquisition

```
Input: stream of sEMG samples x
init model to zero;
foreach hand gesture g do
    while participant performs g do
        acquire new sample x;
        update model with {x, label(g)};
    end
end
```

Algorithm 2: Feedback-Aided Data Acquisition

```
Input: stream of sEMG samples x
init model to zero;
foreach hand gesture g do
    while participant performs g do
        acquire new sample x;
        compute prediction error;
        generate acoustic feedback;
        update model with {x, label(g)};
    end
end
```

Algorithm 3: Feedback-Aided Data Acquisition with Sample Selection

```
Input: stream of sEMG samples x
init model to zero;
foreach hand gesture g do
    while participant performs g do
        acquire new sample x;
        compute prediction error;
        generate acoustic feedback;
        if predictionerror > threshold then
            update model with {x, label(g)};
        end
    end
end
```

The experimental protocol required the completion of all the tasks. If an object was dropped during a manipulation task, the experimenter brought it back to the place where it had been grasped, and the task continued from where it failed. The time needed to reset the position of the object was excluded from the final evaluation of performance. If repeated instabilities of the prosthesis hindered the execution of one task, the experimenter or the participant could suspend the task and request an additional model update. On-demand model updates were obtained with shorter versions of the data acquisition procedure performed at the beginning of the corresponding experimental session. Participants were instructed to hold the malfunctioning hand gesture while randomly moving the arm for 10 s in the area where the task failed, possibly enforcing movements of the shoulder, elbow, and forearm.

2.6. Performance evaluation

We evaluated the effectiveness of each data acquisition procedure based on the duration of the third repetition of the task sequence. We chose the task execution time as an objective measure of performance because it is at the base of many clinical assessment protocols for the hand function and prosthetic control [31]. A limitation of this metric is that, despite being related to the hand functionality, it does not provide information about the movements' quality. Unlike more advanced performance metrics, however, it does not require constraining the experimental setup (e.g., assessing performance in VR) and protocol (e.g., defining Fitt's law style tests), and it does not necessitate evaluation from certified examiners [28].

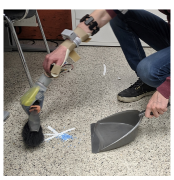
We complemented the task duration with subjective measures of the system's controllability and task difficulty collected in a questionnaire at the end of the experiment. The controllability of the prosthetic system resulting from each training condition was reported on a visual analog scale (VAS) ranging from "very easy to control" to "very difficult to control". Similarly, each task's difficulty was quantified on a VAS ranging from "very difficult" to "very easy". We verified if any of the acquisition strategies resulted in better controllability or faster task execution compared to the others, which could indicate a more robust myocontrol model and, therefore, better training data.

A Shapiro-Wilk test revealed that the task duration and the results of the questionnaire were not normally distributed across participants. For this reason, we used a Friedman test to identify differences in the average value of the statistics of the three training conditions. When the test indicated significant differences, we used repeated post-hoc Wilcoxon signed-rank tests to compare pairs of conditions. We set the significance level of all the tests to $\alpha = 0.05$, and we controlled the inflation of the significance level during repeated pairwise tests by operating a Bonferroni adjustment of the p-value [32]. In this paper, we reported unadjusted p-values (p) for the Friedman tests and Bonferroni-adjusted p-values ($\hat{p}$) for the post-hoc pairwise tests.

3. Results

The performance of the myocontrol model was measured by the duration of the tasks during the third

Table 1. Detailed description of the tasks in each performance evaluation repetition.

Task	Name	Description
	Pour water	A bottle and a jar are placed, respectively, on the lower shelf and on the table. Grasp the bottle^{p}, unscrew the cap^{s}, place the bottle and the cap on the table. Take the jar^{s}, unscrew the lid^{p}, and put it on the table. Take the bottle^{p} and pour the content into the jar. Close the jar^{p} and put it on the table. Take the bottle^{p}, close it^{s}, and bring it back to the lower shelf.
	Serve food	A pot, a plate containing three tennis balls, and a spoon are laid on the table. Use the spoon^{p} to bring the balls from the plate to the pan. Grab the pot by the handle łand tilt it by about 80 degrees, scoop the balls from the pot to the plate using the spoon^{p}.
	Phone and rolling ball	A telephone is on the table. Dial^{p} a sequence of numbers on the phone (1 to 9, 9 to 1, 0, "dial") with an index pointing gesture. A small ball is on the floor, and a target position is marked on the floor about one meter away. Use the index pointing gesture to push the ball^{p} to the target position.
	Pegboard	Three wooden shapes from one pegboard game are laid on the lower shelf, while the base is laid on the higher shelf. Pick^{p} each shape and stack it to the corresponding peg.
	Sweep the floor	A hand broom and a dustpan are placed on the lower shelf, while a bowl and some gravels are laid on the floor. Grab hand broom^{p} and dustpan^{s}, sweep the gravels onto the dustpan, empty the dustpan in the bowl, and bring the hand broom and the dustpan back to the lower shelf.

^{p} prosthetic hand; ^{s} sound hand.

repetition of the task sequence, i.e., the performance evaluation session. Figure 2A reports the duration of the evaluation session corresponding to the three data acquisition strategies. A Friedman test, followed by pairwise post-hoc Wilcoxon tests, revealed that the evaluation session in the FA-DA condition was significantly faster than in the NF-DA condition (average tasks sequence duration of 166.0 s versus 198 s, $W = 19.5$, $\hat{p} = 0.012$). The average duration of the task sequence in the FASS-DA condition, 183 s, did not differ significantly from those of the other conditions. Figure 2B and Figure 2C report the performance of the twelve naive and six experienced participants. For every training strategy, experienced participants completed the evaluation session faster than naive participants. Although not supported by statistical evidence, both groups seemed to perform better after FA-DA compared to NF-DA. The use of feedback during data acquisition reduced the average duration of the performance evaluation session by 15 % for naive participants and by 19 % for experienced participants. For both groups, the mean duration of the tasks after FASS-DA was characterized by high variability, and its average value was between those of the other two conditions. Figure 2D describes the performance of the participants during the individual tasks. Friedman tests were performed for each task and confirmed significant differences in completion time for the third task ($\chi^2(2) = 7.4$, $p = 0.024$). Post-hoc tests, however, failed to identify differences between any pair of conditions, which could be caused by the application of a conservative Bonferroni adjustment to the p-value. Nonetheless, the average duration of every task after FA-DA was slightly lower than after NF-DA.

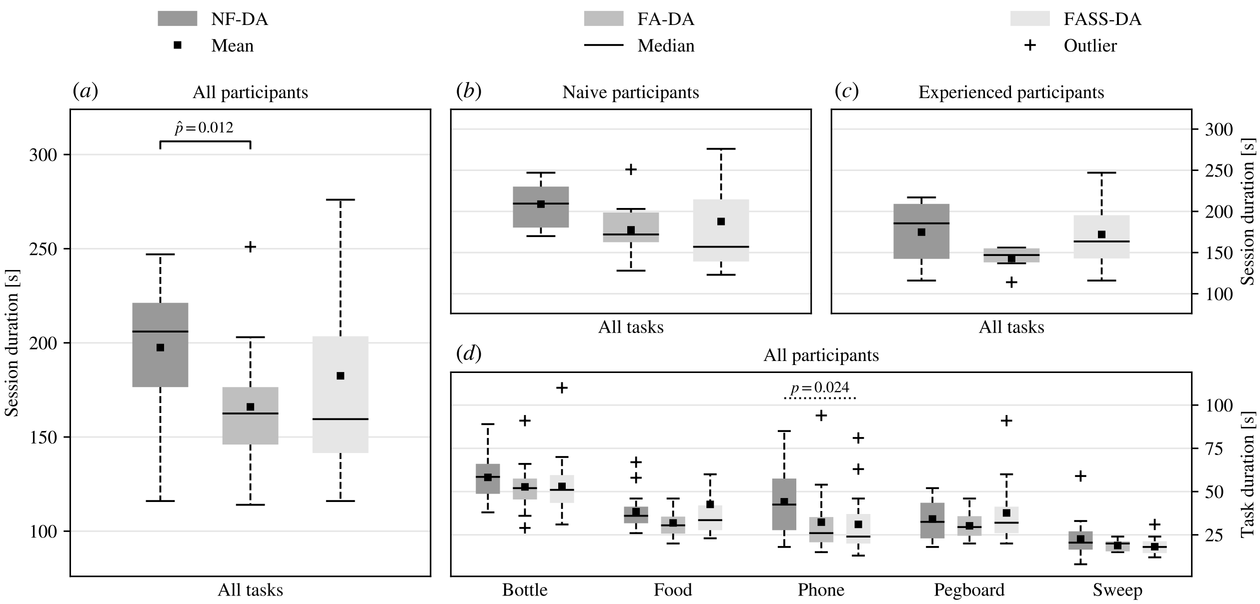


Figure 2. *Duration of the tasks in the performance evaluation session.* *(a)* Participants completed the evaluation session significantly faster in the FA-DA condition compared to the NF-DA condition ($\hat{p}$ Bonferroni-adjusted). The performance in the FASS-DA condition did not differ significantly from those of the other conditions. The same could be observed by either considering the twelve naive *(b)* or the six experienced *(c)* participants. *(d)* The average duration of each task in the FA-DA condition was slightly lower than that measured after NF-DA, although multiple Friedman tests identified significant differences (p unadjusted) only in the duration of the third task (dialing a phone number). In the paper, boxplots' whiskers extend to the most extreme samples within the first quartile -1.5 IQR and the third quartile $+1.5$ IQR.

The performances of FASS-DA remained equivalent to those of the other strategies.

Figure 3 shows the duration of the three repetitions of the task sequence for naive and experienced participants. We referred to these repetitions as the first and second familiarization sessions (F1 and F2), and the performance evaluation session (E). The participants tested the data collection strategies in randomized orders so to counterbalance possible transfer learning effects. Therefore, the results displayed in the figure follow a chronological order within each training condition but not across different conditions. For all the training conditions, naive participants completed the performance evaluation session around 24 % faster than the first familiarization session, Figure 3A. Friedman tests confirmed that the reduction of the task completion time during the three repetitions was significant in every training condition ($\chi^2(2) = 9.5$, $p = 0.009$ for NF-DA; $\chi^2(2) = 18.2$, $p < 0.001$ for FASS-DA; $\chi^2(2) = 18$, $p < 0.001$ for FASS-DA). At the same time, the variability of the results reduced during the familiarization process. The IQR shrunk from 208-302.5 s to 181.5-229 s for NF-DA, from 191.8-282 s to 163.8-197.5 s for FA-DA, and from 152-327.3 s to 140.3-213.5 s for FASS-DA. These results, taken together, indicate that a strong learning effect took place for naive participants during the familiarization process of each strategy. This learning trend was not as evident among the six experienced participants, Figure 3B. For them, the reduction of the task sequence duration due to familiarization was statistically significant only in the FASS-DA condition ($\chi^2(2) = 9.3$, $p = 0.009$). Nonetheless, the task execution time reduced by approximately 19 % during the familiarization process for all the training strategies.

Table 2. Median amount of training samples acquired and used to build the myocontrol model

Acquisition protocol	**# training samples**
NF-DA	27284.5 (IQR 25993-30696)^{a,b}
FA-DA	28439.5 (IQR 25620-30440)^{a,b}
FASS-DA	26879 (IQR 26654-32696)^{a} 7228.5 (IQR 5345-8646)^{b}

^{a}acquired; ^{b}used.

Table 2 details the median number of training samples acquired by each strategy and the number of samples selected to train the myocontrol model. The number of training samples comprised the data

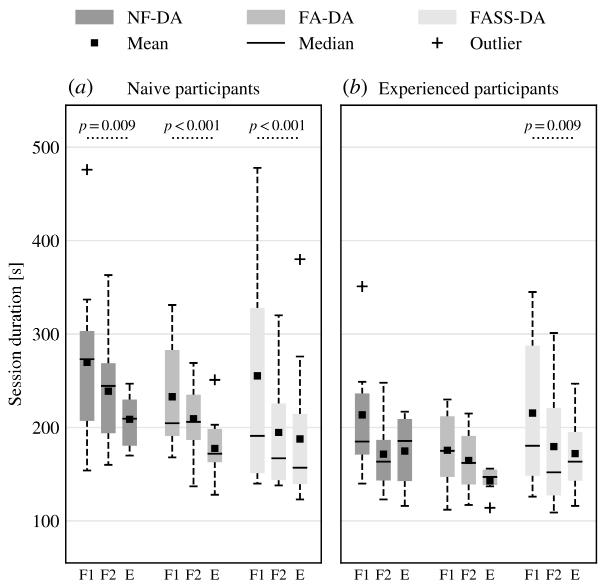


Figure 3. *Effect of learning on the duration of the task sequence.* The three repetitions of the task sequence were labeled F1, first familiarization, F2, second familiarization, and E, performance evaluation session. *(a)* The twelve naive participants showed a significant reduction in the average task completion time due to familiarization (p unadjusted). *(b)* For the six experienced participants, the familiarization with the system significantly reduced the duration of the task session only in the FASS-DA condition.

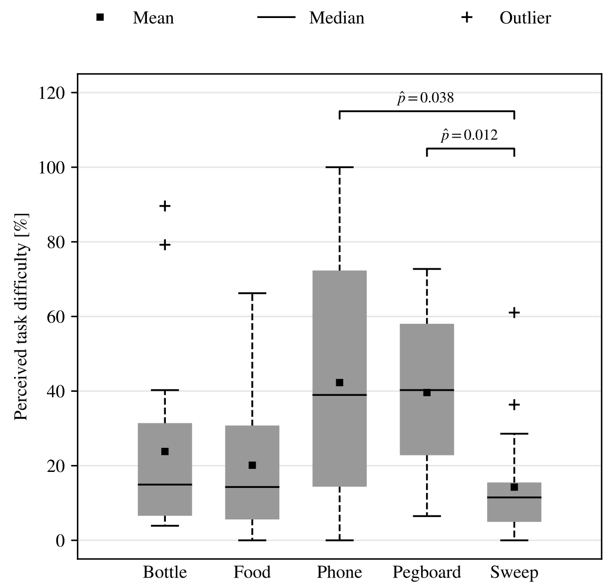


Figure 4. *Perceived tasks difficulty.* On average, participants found more difficult those tasks that required to manipulate small objects (completing the pegboard) or to precisely touch small target areas (dialing a phone number). This result was only partially supported by statistical evidence ($\hat{p}$ Bonferroni-adjusted).

acquired during the initial acquisition and during all the on-demand model updates requested by the participants. All the strategies acquired a comparable amount of training samples, about 28000, although with some variations. The median number of acquired training samples was approximately 27000 for NF-DA and FASS-DA, and 28500 for FA-DA. The median number of on-demand model updates was equal to 0.5 (IQR 0-2) for NF-DA, 1 (IQR 0-2) for FA-DA, and 0 (IQR 0-3) for FASS-DA. While NF-DA and FA-DA used all the training samples to build the myocontrol mode, FASS-DA only employed a median of $\approx$ 7000 samples, roughly corresponding to a quarter of the acquired data.

Figure 4 shows the perceived difficulty of the myocontrol tasks, assessed by the participants in the final questionnaire, and converted into a percentage from 0 % ("very easy") to 100 % ("very difficult"). The ratings seemed to split tasks into two groups. The relative difficulty of pouring water, serving food, and sweeping the floor was relatively low, around 20 % on average. Precision tasks such as dialing a phone number and completing a pegboard were given a higher average difficulty, around 40 %. In particular, a quarter of the participants found it extremely difficult to dial phone numbers, as they reported a difficulty level higher than 75 %, more than what was reported for all the other tasks. A Friedman test confirmed the existence of significant differences in the perceived complexity of the tasks ($\chi^2(4) = 23.1$, $p < 0.001$). Pairwise post-hoc tests, however, only confirmed that the sweeping task was easier than the dialing task ($W = 9$, $\hat{p} = 0.004$) and the pegboard task ($W = 11$, $\hat{p} = 0.0012$).

The controllability of the prosthetic hand during the myocontrol tasks, reported by the participants in the questionnaire, was converted into a percentage from 0 % ("very difficult to control") to 100 % ("very easy to control"). Overall, the use of feedback during the data acquisition resulted in an improvement of the controllability level of about 10% compared to acquiring data without feedback (controllability level of 60 % for NF-DA, 70 % for FA-DA, 71 % for FASS-DA), Figure 5A. A Friedman test, however, did not support this finding with statistical evidence ($\chi^2(2) = 51$, $p = 0.19$). Naive participants reported lower controllability for every training condition, by about 22% on average, compared to experienced participants. In any training condition, the average controllability reported by naive participants was about 22 % lower than that reported by experienced participants. The ratings of naive participants were mixed. Although the controllability level was slightly higher for the feedback-aided acquisition strategies (controllability

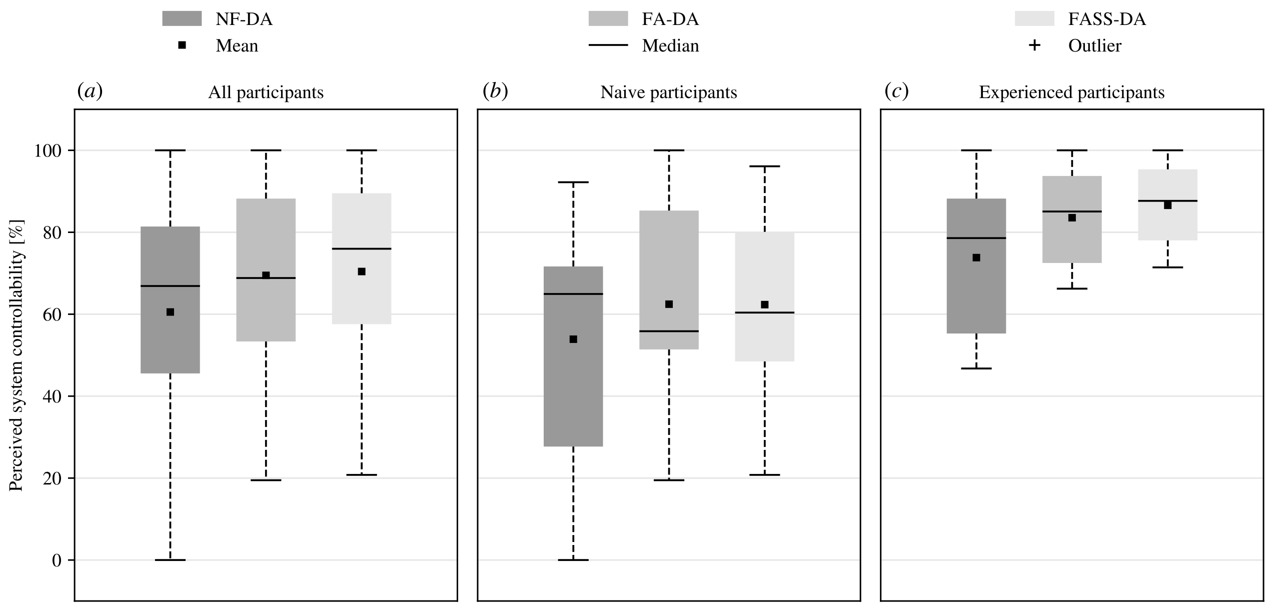


Figure 5. *Perceived controllability of the prosthetic hand. (a)* On average, participants found that the prosthetic system was easier to control after each of the HL data acquisitions compared to the OL data acquisition. *(b)* The ratings reported by the twelve naive participants were mixed and, on average, lower than those of the experienced participants. This caused the high variability observed in the overall results and possibly explained the lack of statistical significance. *(c)* The six experienced participants consistently reported that data acquisition routines with feedback resulted in better controllability of the prosthetic system.

level of 54 % for NF-DA and 62 % for FA-DA and FASS-DA), the spread of the ratings was exceptionally high, especially for the data acquisition without feedback (interquartile range, IQR, equal to 28-71 %). Experienced participants, conversely, reported sharper improvements in controllability by following feedback-aided training strategies. The average controllability increased from 74 % for NF-DA to 84 % for FA-DA, and 87 % for FASS-DA. The spread of these results was lower than that observed in naive participants (IQR equal to 56-88 % for NF-DA, 73-94 % for FA-DA, and 78-95 % for FASS-DA). However, this result was not supported by statistical evidence, possibly due to the limited amount of experienced participants.

4. Discussion and conclusions

We implemented a feedback-aided data acquisition and model building protocol in which the myocontrol model is trained online, while the participant receives instantaneous auditory feedback about the usefulness of the recorded training samples. In the experiment, we have compared two variants of the feedback-aided acquisition strategy to a traditional one, in which the user performs the acquisition routine without automatic guidance. Our results confirm that *automatically guiding the user during the data acquisition yields better myocontrol*, both *objectively*, enabling faster completion of tasks and requiring less computation space and power, and *subjectively*, increasing the perceived controllability of the prosthesis reported in quantitative questionnaires.

Participants completed the sequence of manipulation tasks significantly faster when using FA-DA compared to NF-DA, Figure 2. The average task duration after FA-DA was about 16 % shorter than after NF-DA. The performance offered by FASS-DA was characterized by higher variability and did not differ significantly from those of the other acquisition strategies. Even though half of the participants performed equivalently well with FASS-DA and FA-DA, the other half showed considerably worse performance for FASS-DA.

This may indicate that the sample selection criterion used in FASS-DA was too strict (the system discarded all the training samples that determined a prediction error below 5 % of the maximum prediction error). By relaxing that criterion, the performance of FASS-DA should tend to those of FA-DA, therefore, at least, reducing the variability. Careful tuning of the sample selection criterion should be considered for future investigation.

Nonetheless, FASS-DA considerably reduced the number of samples used to train the machine, of about three-quarters of the total on average. This is especially relevant for realtime applications where the myocontrol model needs to be updated incrementally, requiring repeated model updates for batches of incoming training samples. We then conclude that FASS-DA can be used as a second choice over FA-DA, only when less computational resources are available to the machine learning system.

Feedback-aided acquisition improved myocontrol performance both for the 6 experienced and 12 naive participants (after a short familiarization with the system), as it is apparent from Figure 2B and Figure 2C. For experienced participants, in particular, FA-DA reduced the average task duration by almost 20 % compared to NF-DA.

This suggests that even experienced myocontrol users could benefit from automatic guidance to identify the model's weaknesses during data acquisition. Figure 2D, finally, suggests that feedback-aided data acquisition improved myocontrol performance uniformly for each task.

As it was predictable, a quite evident learning effect is present in all participants' performance, from the two familiarization phases to the experimental one, Figure 3. Interestingly, this trend characterized both naive and experienced participants, albeit less so in the latter case. On the one hand, this means that the effect of feedback can be observed after a short familiarization with the system (by inexperienced participants, that is). On the other hand, automatic guidance during data acquisition retains its usefulness over time, since it allows experienced users to identify and address the flaws of the myocontrol system already during the data acquisition.

During the familiarization, participants learn to compensate distracting factors that are inherent in the myocontrol of the prosthetic device, such as the latency and the weight of the robotic hand, and the non-intuitive control of the contraction strength (nonlinear algorithms may not guarantee monotonic mappings of muscle contraction to grip strength). This contributes to reducing the variability of the results and, therefore, helps observe the effects of interest, such as the effect of different data acquisition procedures. We observed that during the familiarization process, the performance of naive participants decreased in variability and seemed to tend to those of experienced participants. However, the average duration of the task session at the end of the familiarization process remained slightly higher for naive participants. This might indicate that their performance could have further improved with a longer familiarization.

Data acquisition with feedback improved the perceived controllability of the prosthesis by about 10 % on average, Figure 5. The average controllability reported after FA-DA was similar to FASS-DA and higher than NF-DA. However, this difference was not statistically significant due to the large variability in the results. Naive participants provided disparate opinions regarding the system's controllability, and a gap of about 20 % divided the average controllability reported by naives and experienced participants. More focused questions could have been beneficial to reduce this variance. Nonetheless, the improvement reported by experienced participants exhibited a distinct trend in favor of the acquisition strategies with feedback.

The tasks perceived as most difficult were dialing a phone number and completing the pegboard game, Figure 4. They involved touching small target areas with the index finger's tip and placing small objects in positions difficult to reach. This required a significative amount of arm movements to compensate for the lack of an active wrist, which elicited the limb position effect. Interestingly, the only task in which the improvement between NF-DA and the other strategies was statistically significant was dialing the phone number, Figure 2D. This seems to indicate that the proposed feedback guidance improved the model's robustness, especially in tasks mostly affected by the limb position effect.

Interactivity during the training and during the use of the prosthesis lets the user develop more trust in the prosthesis through the usage of a friendly interface, dexterous but straightforward at the same time. In [9], a partially satisfactory result appears, mainly due, we speculate, to a suboptimally designed interaction protocol. All in all, however, the benefits of feedback guidance are not guaranteed to transfer to disabled users, and this issue must be further investigated. We have dealt with this problem already in Gigli et al. [23], to which we refer the interested reader.

Our experiment focused on evaluating the effectiveness of automatic feedback guidance in identifying and counteracting the limb position effect. To be used in everyday prosthetics, this methodology must be adapted and validated against other sources of variability in the myoelectric signal, such as electrodes shift and pressure changes within the socket due to the prosthesis weight. However, we notice that the proposed feedback targets the model's mispredictions regardless of their cause. For this reason, we argue that this feedback may be used with limited adaptation in less constrained settings.

The proposed feedback uses the prediction error to identify when the model's predictions are negatively affected by the limb position. This criterion assumes that the user is able to perform and maintain the correct gesture during the data acquisition. This can be guaranteed for non-disabled users thanks to proprioceptive and visual feedback of their hand configuration, but not for amputees. A transradial amputee, particularly a naive or distracted one (in distracting conditions), might struggle to maintain a consistent muscular activation during the data acquisition, triggering the acoustic feedback for arm configurations where more data is not required. That circumstance is potentially disruptive for the acquisition algorithm since it may lead the user to acquire more training samples while providing a wrong muscular pattern or in arm configurations that do not need to be reinforced with more data. Users with higher amputation levels could struggle even more since they might also have problems following the proper arm trajectory. This problem might be mitigated by inducing the user to perform more consistent and repeatable muscular activations. One could do so by exploiting bilateral mirrored training or using the prosthetic hand as a proxy for the missing limb during the data acquisition [33]. If this sort of training proves ineffective, then the feedback should be redesigned based on other metrics. Competences ranging from psychology to human-machine interfaces design, as well as focus groups and user studies, will be required to solve this problem.

4.1. Conclusion

This work shows that providing automatic feedback guidance during training data acquisition can improve the robustness of myocontrol models to the limb position effect. Data acquisition for myocontrol is often conducted without providing feedback guidance to the user, which may undermine the quality of the recorded data. We implemented a novel data acquisition protocol that collects myoelectric signals dynamically in multiple arm positions while building the model online and guiding the user with instantaneous acoustic feedback. We designed the feedback to induce the user to hover with the arm in the areas of the peripersonal space characterized by poor intent detection, i.e., a discrepancy between the model's prediction and the ground truth. In our experiment, data acquisition strategies that guided the participant to identify and acquire more training samples in problematic areas of the input space yielded better performance, both objective and subjective, and granted the participant a better understanding of the system.

Acknowledgments

This work was partially supported by the DFG project Deep-Hand, CA1389/1-2.

References

[1] Micera S, Carpaneto J and Raspopovic S 2010 "Control of hand prostheses using peripheral information" *IEEE reviews in biomedical engineering* **3** 48–68

[2] Cordella F, Ciancio A L, Sacchetti R, Davalli A, Cutti A G, Guglielmelli E and Zollo L 2016 "Literature review on needs of upper limb prosthesis users" *Frontiers in neuroscience* **10** 209

[3] Resnik L, Ekerholm S, Borgia M and Clark M A 2019 "A national study of veterans with major upper limb amputation: Survey methods, participants, and summary findings" *PloS one* **14** e0213578

[4] Reiber G E, McFarland L V, Hubbard S, Maynard C, Blough D K, Gambel J M and Smith D G 2010 "Servicemembers and veterans with major traumatic limb loss from vietnam war and oif/oef conflicts: Survey methods, participants, and summary findings" *Journal of rehabilitation research and development* **47** 275–298

[5] Atkins D J, Heard D C and Donovan W H 1996 "Epidemiologic overview of individuals with upper-limb loss and their reported research priorities" *Jpo: Journal of prosthetics and orthotics* **8** 2–11

[6] Hahne J M, Biessmann F, Jiang N, Rehbaum H, Farina D, Meinecke F C, Müller K R and Parra L C 2014 "Linear and nonlinear regression techniques for simultaneous and proportional myoelectric control" *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **22** 269–279

[7] Parajuli N, Sreenivasan N, Bifulco P, Cesarelli M, Savino S, Niola V, Esposito D, Hamilton T J, Naik G R, Gunawardana U *et al.* 2019 "Real-time emg based pattern recognition control for hand prostheses: a review on existing methods, challenges and future implementation" *Sensors* **19** 4596

[8] Sensinger J W, Lock B A and Kuiken T A 2009 "Adaptive pattern recognition of myoelectric signals: exploration of conceptual framework and practical algorithms" *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **17** 270–278

[9] Nowak M, Castellini C and Massironi C 2018 "Applying Radical Constructivism to machine learning: a pilot study in assistive robotics" *Constructivist Foundations* **13** 250–262 URL `http://constructivist.info/13/2/250.nowak`

[10] Meattini R, Nowak M, Melchiorri C and Castellini C 2019 "Automated instability detection for interactive myocontrol of prosthetic hands" *Frontiers in Neurorobotics* **13** URL `https://www.frontiersin.org/articles/10.3389/fnbot.2019.00068`

[11] Meattini R, Biagiotti L, Palli G and Melchiorri C 2019 "Grasp-oriented myoelectric interfaces for robotic hands: A minimal-training synergy-based framework for intent detection, control and perception" *International Workshop on Human-Friendly Robotics* (Springer) pp 110–124

[12] Hahne J M, Markovic M and Farina D 2017 "User adaptation in myoelectric man-machine interfaces" *Scientific Reports* **7** 4437

[13] Shehata A W, Scheme E J and Sensinger J W 2018 "Improving internal model strength and performance using augmented feedback" *bioRxiv* 259754

[14] Marasco P D, Kim K, Colgate J E, Peshkin M A and Kuiken T A 2011 "Robotic touch shifts perception of embodiment to a prosthesis in targeted reinnervation amputees" *Brain* **134** 747–758

[15] Woodward R B and Hargrove L J 2019 "Adapting myoelectric control in real-time using a virtual environment" *Journal of neuroengineering and rehabilitation* **16** 11

[16] Hahne J M, Dähne S, Hwang H J, Müller K R and Parra L C 2015 "Concurrent adaptation of human and machine improves simultaneous and proportional myoelectric control" *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **23** 618–627

[17] Campbell E, Phinyomark A and Scheme E 2020 "Current trends and confounding factors in myoelectric control: Limb position and contraction intensity" *Sensors* **20**

1613

[18] Khushaba R N, Takruri M, Miro J V and Kodagoda S 2014 "Towards limb position invariant myoelectric pattern recognition using time-dependent spectral features" *Neural Networks* **55** 42–58

[19] Betthauser J L, Hunt C L, Osborn L E, Masters M R, Lévay G, Kaliki R R and Thakor N V 2018 "Limb position tolerant pattern recognition for myoelectric prosthesis control with adaptive sparse representations from extreme learning" *IEEE Transactions on Biomedical Engineering* **65** 770–778

[20] Fougner A, Scheme E, Chan A D, Englehart K and Stavdahl Ø 2011 "Resolving the limb position effect in myoelectric pattern recognition" *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **19** 644–651

[21] Scheme E, Biron K and Englehart K 2011 "Improving myoelectric pattern recognition positional robustness using advanced training protocols" *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (IEEE) pp 4828–4831

[22] Radmand A, Scheme E and Englehart K 2014 "On the suitability of integrating accelerometry data with electromyography signals for resolving the effect of changes in limb position during dynamic limb movement" *JPO: Journal of Prosthetics and Orthotics* **26** 185–193

[23] Gigli A, Gijsberts A and Castellini C 2020 "The merits of dynamic data acquisition for realistic myocontrol" *Frontiers in Bioengineering and Biotechnology*

[24] Sierra González D and Castellini C 2013 "A realistic implementation of ultrasound imaging as a human-machine interface for upper-limb amputees" *Frontiers in Neurorobotics* **7** URL `https://www.frontiersin.org/articles/10.3389/fnbot.2013.00017`

[25] Gijsberts A, Bohra R, Sierra González D, Werner A, Nowak M, Caputo B, Roa M A and Castellini C 2014 "Stable myoelectric control of a hand prosthesis using non-linear incremental learning" *Frontiers in Neurorobotics* **8** URL `https://www.frontiersin.org/articles/10.3389/fnbot.2014.00008`

[26] Patel G K, Nowak M and Castellini C 2017 "Exploiting knowledge composition to improve real-life hand prosthetic control" *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **25** 967–975

[27] Strazzulla I, Nowak M, Controzzi M, Cipriani C and Castellini C 2017 "Online bimanual manipulation using surface electromyography and incremental learning" *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **25** 227–234 URL `https://ieeexplore.ieee.org/document/7460959`

[28] Wang S, Hsu J, Trent L, Ryan T, Kearns N, Civillico E and Kontson K 2018 "Evaluation of performance-based outcome measures for the upper limb: a comprehensive narrative review" *PM&R*

[29] Hermansson L M, Fisher A G, Bernspång B and Eliasson A C 2005 "Assessment of capacity for myoelectric control: a new Rasch-built measure of prosthetic hand control" *Journal of rehabilitation medicine* **37** 166–71

[30] Kyberd P J, Murgia A, Gasson M, Tjerks T, Metcalf C, Chappell P H, Warwick K, Lawson S E and Barnhill T 2009 "Case studies to demonstrate the range of applications of the Southampton Hand Assessment Procedure" *British Journal of Occupational Therapy* **72** 212–218

[31] Vujaklija I, Roche A D, Hasenoehrl T, Sturma A, Amsuess S, Farina D and Aszmann O C 2017 "Translating research on myoelectric control into clinics—are the performance assessment methods adequate?" *Frontiers in neurorobotics* **11** 7

[32] Chen S Y, Feng Z and Yi X 2017 "A general introduction to adjustment for multiple comparisons" *Journal of thoracic disease* **9** 1725

[33] Igual C, Pardo L A, Hahne J M and Igual J 2019 "Myoelectric control for upper limb prostheses" *MDPI Electronics* **8** 1244